

Interaction between Syntax Processing in Language and in Music: An ERP Study

Stefan Koelsch^{1,2}, Thomas C. Gunter¹, Matthias Wittfoth³,
and Daniela Sammler¹

Abstract

■ The present study investigated simultaneous processing of language and music using visually presented sentences and auditorily presented chord sequences. Music-syntactically regular and irregular chord functions were presented synchronously with syntactically correct or incorrect words, or with words that had either a high or a low semantic cloze probability. Music-syntactically irregular chords elicited an early right anterior negativity (ERAN). Syntactically incorrect words elicited a left anterior negativity (LAN). The LAN was clearly reduced when words were presented simultaneously with music-syntactically irregular chord functions. Processing

of high and low cloze-probability words as indexed by the N400 was not affected by the presentation of irregular chord functions. In a control experiment, the LAN was not affected by physically deviant tones that elicited a mismatch negativity (MMN). Results demonstrate that processing of musical syntax (as reflected in the ERAN) interacts with the processing of linguistic syntax (as reflected in the LAN), and that this interaction is not due to a general effect of deviance-related negativities that precede an LAN. Findings thus indicate a strong overlap of neural resources involved in the processing of syntax in language and music. ■

INTRODUCTION

The question of (non)specificity of neural mechanisms that underlie the processing of music and language has appreciated increasing interest in cognitive neuroscience during the past years (Koelsch, Kasper, et al., 2004; Patel, 2003; Koelsch, Gunter, von Cramon, et al., 2002; Maess, Koelsch, Gunter, & Friederici, 2001; Zatorre & Peretz, 2001; Besson, Faita, Peretz, Bonnel, & Requin, 1998; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). A study from Patel et al. (1998) compared structural processing in music and language. Results showed that the processing of structural incongruities in both domains elicits a P600 that does not differ between domains, presumably because the same neuronal resources are involved in processes of structural integration (Patel, 1998). Based on these findings, Patel (1998, 2003) suggested the *shared syntactic integration resource hypothesis* (SSIRH), which assumes that the overlap in the syntactic processing of language and music can be conceived of as an overlap in the neural areas and operations which provide the resources of syntactic integration.

Similarly, an early left anterior negativity (ELAN), which is taken to reflect initial syntactic structure build-

ing (Friederici, 2002), resembles the early right anterior negativity (ERAN), which is taken to reflect processing of music-syntactic information (Koelsch & Friederici, 2003). The similarity of ELAN and ERAN has also been suggested to be due to overlapping neural resources involved in the processing of syntactic information in language and music. Corroboratingly, the neural generators of the ERAN have been localized in the inferior fronto-lateral cortex (inferior BA 44; Maess et al., 2001), areas that are also involved in the processing of syntactic information during language perception (Friederici, 2002). Similar activations of the inferior fronto-lateral cortex have been reported by functional imaging studies investigating the processing of musical structure (Patel, 2003; Tillmann, Janata, & Bharucha, 2003; Janata et al., 2002; Koelsch, Gunter, von Cramon, et al., 2002; Parsons, 2001; Platel et al., 1997; for a review, see Koelsch, 2005). However, so far there is a lack of studies investigating the simultaneous processing of syntax in language and music (for studies investigating simultaneous processing of structure in music and semantics in language, see Poulin-Charronnat, Bigand, Madurell, & Peereman, 2005; Bonnel, Faita, Peretz, & Besson, 2001; Besson et al., 1998; see also below).

The present study focuses on the question of whether processing of syntax in music interacts with the processing of syntax and semantics in language during the simultaneous processing of music (chords) and language (words). In the language domain, we investigated

¹Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, ²Harvard Medical School, ³University of Bremen, Germany

the left anterior negativity (LAN, e.g., Friederici, 2002) elicited by a (syntactic) gender violation,¹ and the N400 (e.g., Kutas & Federmeier, 2000) elicited by words with a semantic low cloze probability. In the music domain, we investigated the ERAN and the N5. The ERAN can be elicited within harmonic progressions by music-syntactically irregular chords (Koelsch, Schröger, & Gunter, 2002; Maess et al., 2001; Koelsch, Gunter, Schröger, & Friederici, 2000). Usually, the ERAN is followed by a late negativity (N5), which is taken to reflect processes of harmonic integration and is reminiscent of the N400 (Koelsch, Gunter, Schröger, & Friederici, 2003; Koelsch, Gunter, Friederici, et al., 2000).

As linguistic stimuli we used visually presented sentences that were similar to those used in a previous ERP study (Gunter, Friederici, & Schriefers, 2000). Likewise, the auditorily presented chord sequences used in the present study were virtually identical to those of some previous ERP studies (e.g., Koelsch, Grossmann, Gunter, Hahne, & Friederici, 2003; Koelsch, Schmidt, & Kansok, 2002; Koelsch, Schröger, et al., 2002; Koelsch, Gunter, Schröger, Tervaniemi, et al., 2001; Koelsch, Gunter, Friederici, et al., 2000). Here, the mentioned language study (investigating LAN and N400) and previous music studies (investigating ERAN and N5) are combined: five-word sentences were presented visually simultaneously with auditorily presented five-chord sequences. Each word was presented with the onset of a chord (Figure 1).

Corresponding to Gunter, Friederici, et al. (2000), three different sentence types were used: The first type was a syntactically correct sentence in which the final

noun had a high semantic cloze probability. The other two types were modified versions of the first sentence type: (i) a sentence in which the final noun had a low semantic cloze probability, and (ii) a sentence with a gender disagreement between the last word (noun) on the one hand, and the prenominal adjective as well as the definite article that preceded the adjective on the other (see Figure 1).²

Half of the musical sequences ended on a regular tonic chord, the other half ended on a music-syntactically irregular chord function (Neapolitan sixth chord). The irregular chords were major chords that sound normal when presented in isolation, moderately unexpected when presented instead of a subdominant (a Neapolitan may be regarded as a subdominant variation), and strongly unexpected when presented instead of a tonic at the end of a harmonic progression (e.g., Koelsch, Gunter, Friederici, et al., 2000). The regularities that guide the arrangement of chord functions within harmonic progressions have been denoted as part of a musical syntax (Koelsch & Friederici, 2003; Tillmann, Bharucha, & Bigand, 2000; Sloboda, 1985).

Sentences and chord sequences were combined in a 3×2 design (3 sentence types, 2 chord types) so that six different experimental conditions could be investigated: Final nouns of sentences that were syntactically correct and had a high cloze probability were presented simultaneously with either a regular or an irregular final chord function. Analogously, syntactically correct, but semantically unexpected (low cloze probability), final words were presented with either a regular or an irregular chord function ending. Likewise, final words of sentences with a syntactic gender disagreement (and high semantic cloze probability) were presented with either a regular or an irregular chord function (Figure 1). Participants were asked to ignore the musical stimulus, to concentrate on the words, and to answer in 10% of the trials whether the last sentence was correct or (syntactically or semantically) incorrect.

If language processing operates independently of music processing, LAN and N400 should not be influenced by the syntactic irregularities in music (and vice versa). Because both ERAN and N5 can be elicited under ignore conditions (Koelsch, Schröger, et al., 2002; Koelsch, Gunter, Schröger, Tervaniemi, et al., 2001), we hypothesized that, despite the task-irrelevance of chords, irregular chords would elicit an ERAN and an N5. Because of the mentioned overlap of cerebral structures and neuronal processes involved in the syntactic analysis of music and of language, we also hypothesized that the presentation of a music-syntactic irregularity would influence the processing of syntactic violations within the sentences.

No predictions were made about possible interactions between processing of music-syntactic irregularities and semantic language processing. Previous studies investigating this issue with event-related potentials (ERPs)

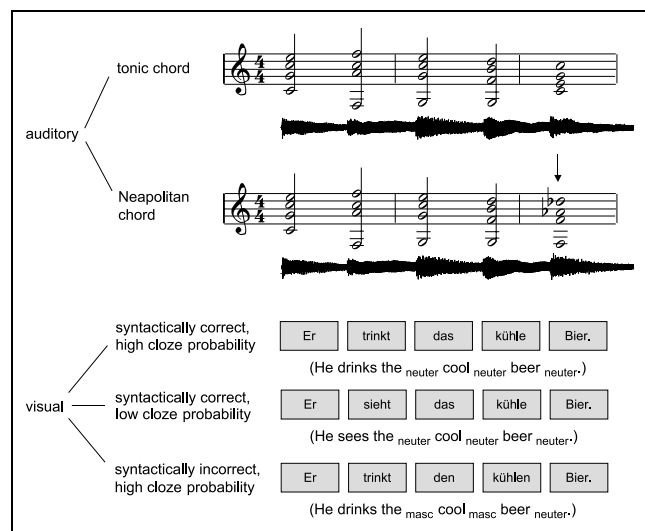


Figure 1. Examples of experimental stimuli. Top: examples of two chord sequences in C major, ending on a regular (upper row) and an irregular chord (lower row, the irregular chord is indicated by the arrow). Bottom: examples of the three different sentence types. Onsets of chords (presented auditorily) and words (presented visually) were synchronous.

(Besson et al., 1998) and behavioral measures (Poulin-Charronnat et al., 2005; Bonnel et al., 2001) do not yet yield a consistent picture: In the studies from Bonnel et al. (2001) and Besson et al. (1998), the occurrence of harmonically regular and irregular notes at the end of a melody did not have an effect on the semantic processing of congruous and incongruous words that were sung on these notes (in the study from Besson et al., 1998, the N400 was used as an electrophysiological index of semantic processing, irregular notes elicited a late positive component). By contrast, a recent behavioral study from Poulin-Charronnat et al. (2005) reports interactions between the processing of structure in music and linguistic semantics using sung chord sequences. The difference between the latter study and the studies from Bonnel et al. and Besson et al. might be due to the different task (a lexical decision task was used in the study from Poulin-Charronnat et al., 2005, vs. explicit congruity judgments in the two other studies), or by the different musical material (chords vs. melodies). The present study is different from the studies of Bonnel et al. and Besson et al. in that chords were used, and is different from the study of Poulin-Charronnat et al. in that chords were presented auditorily and words were presented visually. Thus, no directed hypotheses were made regarding possible influences of music-structural processing on the processing of semantic incongruities.

EXPERIMENT 1

Results

Behavioral responses were evaluated only with respect to the syntax of the sentences, because only the syntactic violations resulted in clear anomalies (the low cloze-probability sentences did not represent semantic violations). Participants scored with 95.5% correct responses (range 81–100%); a *t* test on the percentages of correct responses revealed that participants performed well above chance level [$t(25) = 36.9, p < .0001$].

Regular Words, Irregular Chords

Compared to regular tonic chords, irregular (Neapolitan) chords elicited an ERAN that was maximal around 190 msec, and slightly lateralized to the right (Figure 2). The ERAN showed a polarity inversion at mastoid leads (as expected, e.g., Koelsch, Schröger, et al., 2002). Interestingly, no later negativity (N5, usually maximal around 500–550 msec) was observed in the ERP waveforms of irregular chords. Instead, a tonic late positivity was present, being maximal around 500 msec at centroparietal electrode sites. An analysis of variance (ANOVA) for frontal electrode regions of interest (ROIs, see Methods) for a time interval from 150 to 250 msec with factors Chord-type (regular, irregular) and Hemisphere

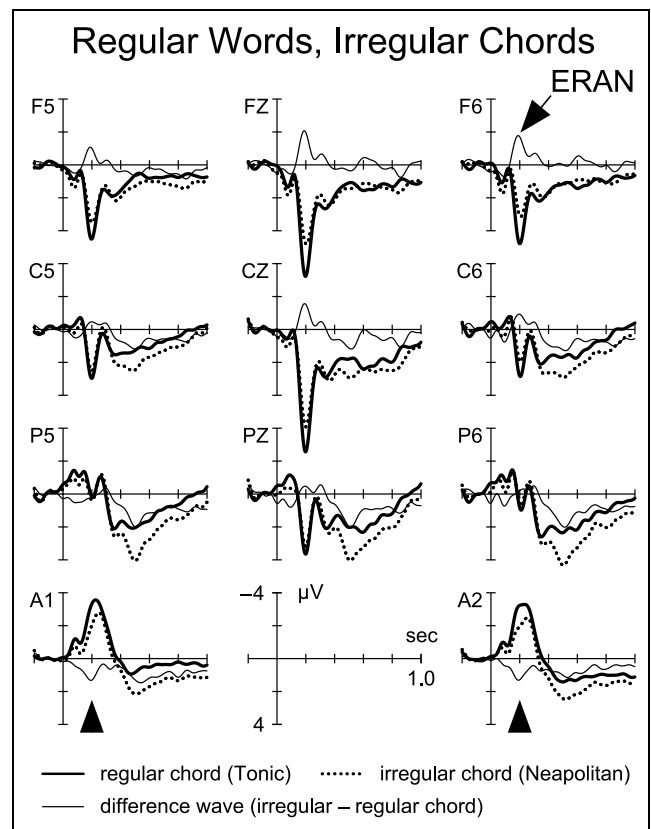


Figure 2. ERPs elicited on regular words (syntactically correct, high cloze probability) when the last chord was a regular tonic chord (solid line), or an irregular Neapolitan chord (dotted line). Irregular chords elicited an ERAN (indicated by the arrow over the thin-lined difference wave). The ERAN inverted polarity at mastoid leads (short arrows in diagrams of A1 and A2, note that a nose-reference was used).

revealed an effect of chord-type [$F(1,25) = 6.83, p < .02$], and an interaction between the two factors [$F(1,25) = 4.23, p < .05$]. The analogous ANOVA for the N5 time interval (350–600 msec) did not indicate an effect of chord-type ($p > .3$).³ An analogous ANOVA for parietal ROIs for the time interval from 450 to 700 msec (late positivity) indicated an effect of chord-type [$F(1,25) = 7.89, p < .01$].

Regular Chords, Irregular Words

Compared to correct words, syntactic (gender) violations elicited a distinct LAN that was maximal around 390 msec (Figure 3A). The LAN was followed by a P600 (as expected, Gunter, Friederici, et al., 2000). An ANOVA for frontal ROIs for a time interval from 300 to 450 msec with factors Syntax (correct, incorrect) and Hemisphere indicated an effect of syntax [$F(1,25) = 12.01, p < .002$], and an interaction between the two factors [$F(1,25) = 23.53, p < .0001$]. The analogous ANOVA for parietal ROIs for a time interval from 450 to 700 msec (P600) also indicated an effect of syntax [$F(1,25) = 19.69, p < .0002$].

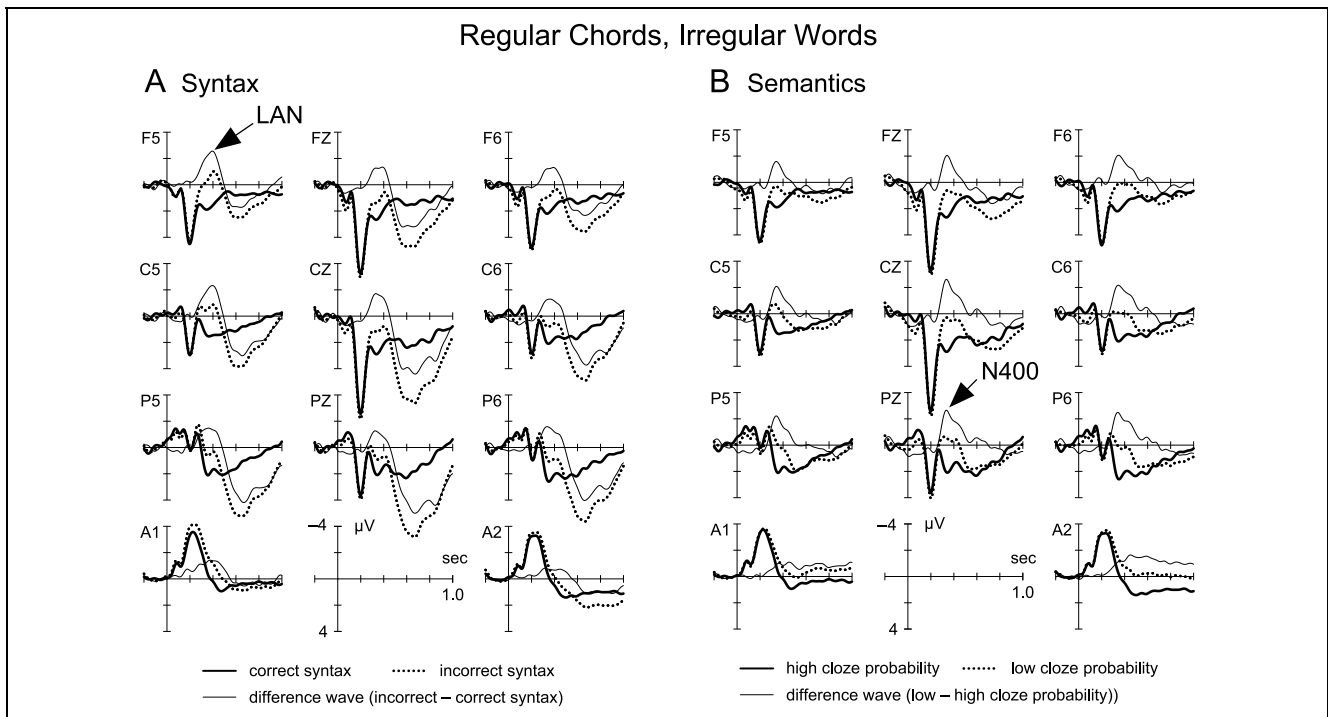


Figure 3. (A) ERPs elicited on regular chords when the last word was syntactically correct (solid line) or syntactically incorrect (dotted line). Syntactically incorrect words elicited an LAN (indicated by the arrow over the thin-lined difference wave; note that both syntactically correct and incorrect words had a high semantic cloze probability). (B) ERPs elicited on regular chords when the last word had a high (solid line), or low cloze probability (dotted line). Semantically unexpected low cloze-probability words elicited an N400 (indicated by the arrow over the thin-lined difference wave; note that both high and low cloze-probability words were syntactically correct, and that, thus, the solid lines of A and B are identical).

Low cloze-probability words (that were semantically less expected) elicited an N400 that was maximal around 350 msec over centro-parietal electrode sites (Figure 3B), and slightly right-lateralized. An ANOVA for parietal ROIs for a 300–450 msec time interval with factors Cloze probability (low, high) and Hemisphere indicated an effect of cloze-probability [$F(1,25) = 12.81, p < .002$], and a two-way interaction [$F(1,25) = 6.81, p < .02$]. To test differences in scalp topography between N400 and LAN, a MANOVA was computed with factors ROIs (four levels: left frontal, right frontal, left parietal, right parietal), Component (LAN, N400), and Condition (correct, incorrect). This MANOVA yielded a three-way interaction [$F(3,75) = 20.41, p < .0001$, Greenhouse–Geisser corrected, epsilon = .71], supporting the observation that LAN and N400 have considerably different scalp topographies.

Syntax × Chords

The former section described the LAN elicited by syntactically incorrect words when sequences ended on regular (tonic) chords. Figure 4 depicts this LAN effect in the solid difference wave. The dotted difference wave of Figure 4 shows, again, the effects of processing syntactically incorrect sentences, but now when words are presented on an irregular chord function. As can be

seen in the two difference waves, the amplitude of the LAN is reduced when words are processed simultaneously with irregular chords.

An ANOVA for frontal ROIs for the LAN time window (300–450 msec) with factors (linguistic) Syntax (correct, incorrect), and Musical regularity (regular, irregular) revealed an effect of syntax [$F(1,25) = 9.58, p < .005$, reflecting that incorrect words elicited an LAN], an effect of musical regularity [$F(1,25) = 6.13, p < .05$, reflecting that ERPs elicited by irregular chords differed from those elicited by regular chords], and an interaction between factors syntax and musical regularity [$F(1,25) = 5.45, p < .03$, reflecting that the LAN effect was smaller when words were presented simultaneously with irregular chords]. The analogous ANOVA for the ERAN time window (150–250 msec) did not reveal an interaction between factors Syntax and Musical regularity [$F(1,25) = 0.55, p > .4$, indicating that the generation of the ERAN was not affected by the syntactic gender violations].

Even when comparing the LAN elicited on regular (tonic) chords with the LAN elicited on irregular (Neapolitan) chords, the LAN amplitude is markedly reduced when elicited on a Neapolitan (Figure 5). That is, even if the LAN could possibly partly overlap with the ERAN, the amplitude of the LAN is smaller when preceded by an ERAN, demonstrating that LAN and ERAN are not additive effects. An ANOVA for frontal ROIs

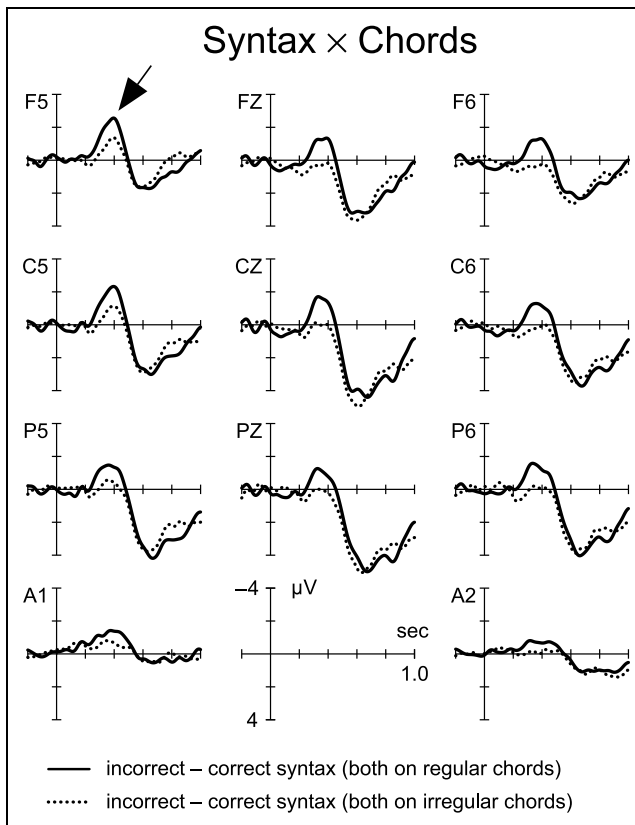


Figure 4. LAN effects (difference waves, syntactically correct subtracted from syntactically incorrect words) for the conditions in which words were presented on regular chords (solid line) and on irregular chords (dotted line). The amplitude of the LAN effect is reduced when syntactically irregular words are processed simultaneously with syntactically irregular chords (arrow).

comparing the ERPs of (a) syntactically incorrect words presented on a regular chord, and (b) syntactically incorrect words presented on an irregular chord revealed a difference between the two conditions [$F(1,25) = 10.18, p < .005$].

Semantics × Chords

In contrast to the LAN, the N400 elicited by low cloze-probability words was not affected when words were presented together with musical irregularities (Figure 6). An ANOVA for parietal ROIs for the N400 time window (350–450 msec) with factors Cloze probability (high, low) and Musical regularity revealed an effect of cloze probability [$F(1,25) = 29.68, p < .0001$, reflecting that low cloze-probability words elicited an N400], and no interaction between factors Cloze probability and Musical regularity ($p > .8$, reflecting that the regularity of chords did not influence the N400 effect). Again, the analogous ANOVA for the ERAN time window (150–250 msec) did not reveal an interaction between the two factors [$F(1,25) = 0.08, p > .7$, reflecting that the generation of the ERAN was not affected by the cloze probability of words].

Discussion

The present data show an interaction between language- and music-syntactic processing: The LAN elicited by syntactically incorrect words was clearly reduced when words were presented simultaneously with music-syntactically irregular chord functions. This effect might be due to overlapping neuronal resources involved in the processing of syntax in both music and language (see General Discussion). However, it is also possible that any kind of deviance-related negativity has an effect on linguistic syntax processing; this issue will be further investigated in Experiment 2.

In contrast, the processing of the semantic aspects of language (indexed by the N400) was not affected by the processing of music-syntactic violations (indexed by the ERAN). This finding is in line with findings of a previous ERP study with sung melodies (Besson et al., 1998) in which the presentation of music-structural irregularities did not have an effect on the processing of semantic aspects of words. However, as mentioned in the Introduction, a behavioral study with sung chord sequences suggested interactions between the processing of musical syntax and (linguistic) semantics (Poulin-Charronnat et al., 2005, in that study, the semantic cloze probability of final words was manipulated, as in the present

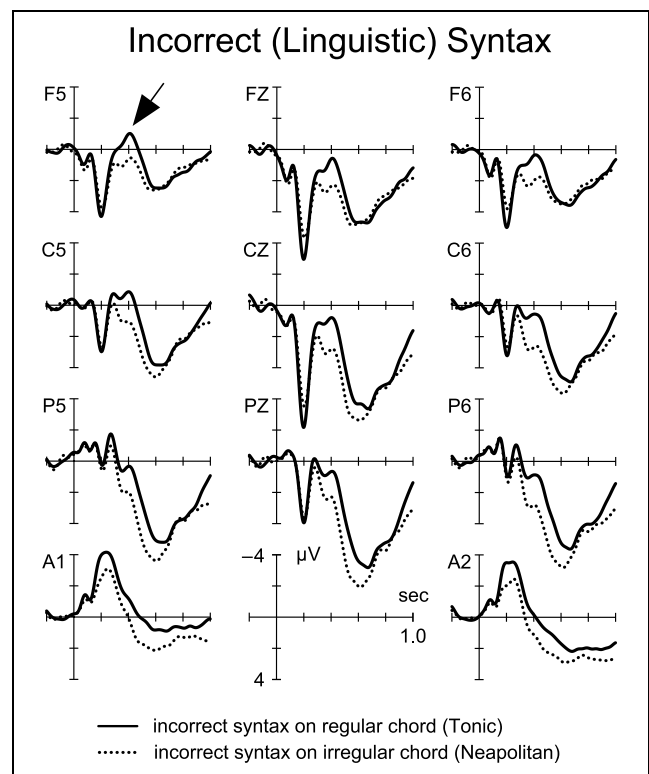


Figure 5. ERPs of syntactically incorrect words presented on a regular chord (solid line), opposed to ERPs of syntactically incorrect words presented on an irregular chord (dotted line). Note the amplitude reduction of the LAN (arrow) when syntactically incorrect words and syntactically incorrect chords are processed simultaneously.

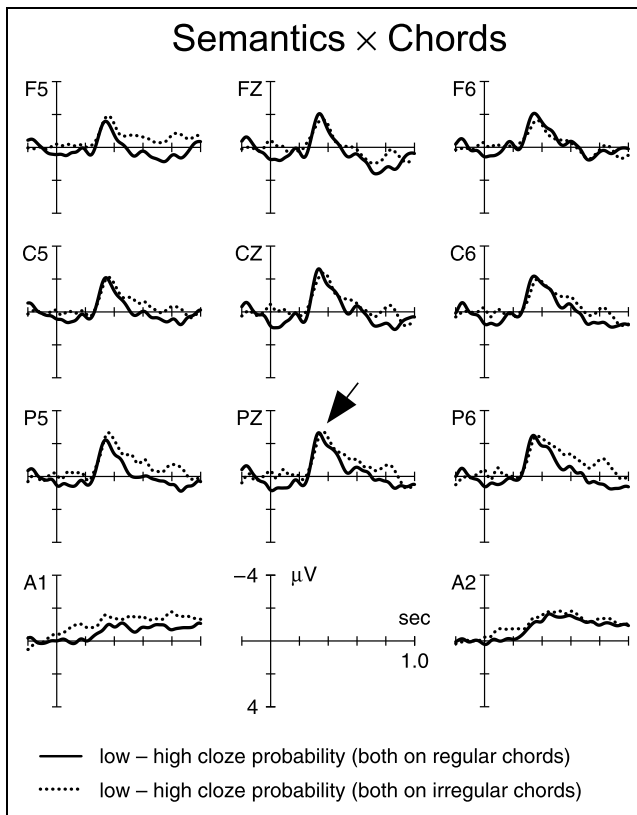


Figure 6. N400 effects (difference waves, high cloze-probability words subtracted from low cloze-probability words) for the conditions in which words were presented on regular chords (solid line) and on irregular chords (dotted line). The amplitude of the N400 was not affected by the simultaneous presentation of irregular chords (arrow).

study). It is possible that the difference between the study of Poulin-Charronnat et al. (2005) and the present study is due to the different stimuli: In the present study, words were presented visually, possibly making it easier to separate the linguistic from the musical information. It is also possible that behavioral measures are more sensitive for detecting interactions between music-syntactic and linguistic-semantic processing. These issues remain to be specified.

Note that in the language domain it is also not yet clear under which conditions syntax and semantics might interact. A study of Gunter, Stowe, and Mulder (1997) did not find an effect of syntax processing (as reflected in the LAN) on semantic processing (as reflected in the N400). On the other hand, a study of Hahne and Friederici (2002) suggests that early processing of syntactic irregularities (reflected in an ELAN) can influence the N400 when participants focus their attention on the syntactic information (in that study, the ELAN was elicited by word category violations). Similarly, Besson and Schön (2001) report attentional influences on the N400 (elicited by semantically incongruous words), as well as on the P600 (elicited by structurally irregular melody notes), during the simultaneous processing of words and

melodies: In that study, the N400 was almost absent when participants ignored the linguistic information and focused their attention on the musical structure.

Taken together, it appears that possible interactions between syntax and semantics in language, as well as in music, have to be specified with respect to (a) the underlying cognitive processes (e.g., processing of phrase structure violations vs. processing of morphosyntactic violations), (b) attentional demands, (c) task, and (d) the type of semantic anomalies used (clear semantic incongruities vs. low cloze probability of words).

Although a clear ERAN was elicited by the music-syntactic irregularities, no N5 was observed, contrary to previous studies that used virtually the same musical stimulus (Koelsch, Grossmann, et al., 2003; Koelsch, Schmidt, et al., 2002; Koelsch, Schröger, et al., 2002; Koelsch, Gunter, Friederici, et al., 2000). Instead, a late positivity was observed. It appears that this late positivity is a P600 (for a review, see, e.g., Friederici, 2004), presumably reflecting processes of syntactic reanalysis after the perception of the music-syntactically irregular chords: Note that this P600 emerged in a condition in which the musical syntax was incorrect, but in which sentences were syntactically *correct* (see Figure 2). That is, participants had to find out that the final word was syntactically correct, although preceding neural activity (elicited by the music-syntactic violations) yielded the presence of a (music-)syntactic irregularity. The P600 presumably reflects these processes of syntactic reanalysis. It is possible that the P600 compensated the negative N5 potentials (which usually emerge in a similar time range). However, it is also possible that focussing on the linguistic information led to a diminution of processes of harmonic integration (and thus, to the absence of the N5), however, this issue remains to be specified.

EXPERIMENT 2

As mentioned above, the data of Experiment 1 do not inform whether the interaction of the processing of linguistic and musical syntax is due to an overlap of the neural processes that mediate both the processing of linguistic and musical syntax, or whether any kind of deviance-related negativity has an effect on linguistic syntax processing.

To address this issue, a second ERP experiment was carried out which was identical to Experiment 1, except that single tones were presented instead of chords. The last tone of a sequence was either a standard tone (in analogy to the regular tonic chord of Experiment 1), or a physically deviant tone (in analogy to the irregular chord). Physically deviant tones presented in a series of standard tones are known to elicit a mismatch negativity (MMN, Schröger, 1998; Näätänen, 1992). An interaction of MMN and LAN would argue against the explanation that ERAN and LAN interact because they both reflect syntactic processes mediated by overlap-

ping neural resources. The absence of an interaction between MMN and LAN would argue in favor of the latter hypothesis, showing that only specific types of auditory irregularity detection (as, e.g., reflected in the ERAN) interact with linguistic syntax processing (as reflected in the LAN).

Results

As in Experiment 1, behavioral responses were evaluated only with respect to the syntax of the sentences. Participants scored with 94.6% correct responses (range 81–100%); a *t* test on the percentages of correct responses revealed that participants performed well above chance level [$t(21) = 35.3, p < .0001$].

Regular Words, Deviant Tones

Compared to standard tones, physically deviant tones elicited an MMN that was maximal around 150 msec over frontal leads (Figure 7). The MMN inverted polarity at mastoidal sites, and was followed by an N2b–P3 complex (the N2b peaked at around 205 msec, the P3 was

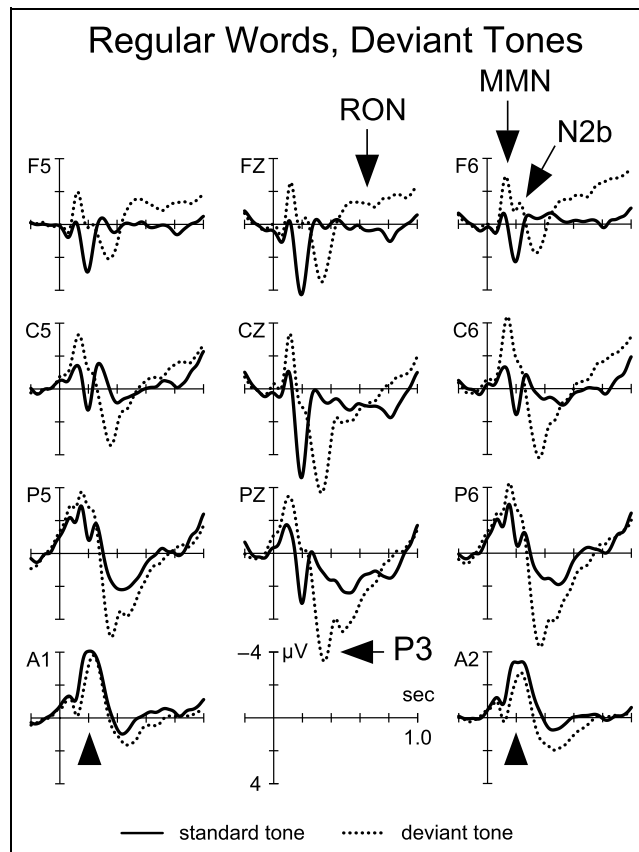


Figure 7. ERPs elicited on regular words (syntactically correct, high cloze probability) when the last tone was a standard (solid line), or a deviant (dotted line). Deviant tones elicited an MMN (long arrow), followed by an N2b–P3 complex, and a RON. The MMN inverted polarity at mastoid leads (short arrows in diagrams of A1 and A2).

maximal at around 345 msec over parietal electrode sites and had a right-hemispheric preponderance). A tonic late negativity emerged around 500 msec and was maximal bilaterally over frontal leads. This late negativity is presumably a reorienting negativity (RON, Schröger & Wolff, 1998), possibly reflecting that participants reoriented their attention back to the linguistic task after being distracted by the physically deviant tones (Schröger, Giard, & Wolff, 2000; Schröger & Wolff, 1998).

An ANOVA for frontal ROIs for a time interval from 90 to 190 msec with factors Tone (standard, deviant) and Hemisphere revealed an effect of tone [$F(1,21) = 23.18, p < .0001$, no two-way interaction]. An analogous ANOVA for parietal ROIs for the P3 time window (300–400 msec) indicated an effect of condition [$F(1,21) = 23.18, p < .0001$], and a two-way interaction [$F(1,21) = 10.42, p < .005$]. An analogous ANOVA for frontal ROIs for the RON time window (600–900 msec) revealed an effect of condition [$F(1,21) = 10.61, p < .005$, no two-way interaction].

Note that in Experiment 1 no late negativity (N5) was present. The difference of the late negative ERPs between the music condition (Experiment 1) and the tone condition (Experiment 2) was significant: A between-subjects ANOVA with factors Stimulus-type (standard, deviant), and Experiment (1, 2) for frontal ROIs revealed a two-way interaction [$F(1,46) = 10.03, p < .005$; time windows used were 350–600 msec [N5 window] for the data from Experiment 1, and 600–900 msec [RON window] for the data from Experiment 2].

Standard Tones, Irregular Words

As in Experiment 1, syntactic violations elicited a distinct LAN that was maximal around 350 msec (Figure 8A). An ANOVA for frontal ROIs for a time interval from 300 to 450 msec with factors Syntax and Hemisphere indicated an effect of syntax [$F(1,21) = 7.72, p < .05$], and an interaction between the two factors [$F(1,21) = 47.36, p < .0001$].

Interestingly, the LAN was more strongly lateralized in Experiment 2 compared to Experiment 1, where the LAN was also lateralized, but distributed more broadly, and also clearly present over the right hemisphere (see Figures 3A and 8A). To test differences in lateralization of the LAN between experiments, a between-subjects ANOVA was computed for frontal ROIs (300–450 msec) with factors Syntax, Hemisphere, and Experiment. This ANOVA indicated a three-way interaction [$F(1,46) = 4.34, p < .05$].

Similarly to Experiment 1, semantically unexpected low cloze-probability words elicited an N400 that was maximal around 400 msec over centro-parietal electrode sites (Figure 8B). An ANOVA for parietal ROIs for a 300–450 msec time interval with factors Cloze probability (low, high) and Hemisphere revealed an effect of

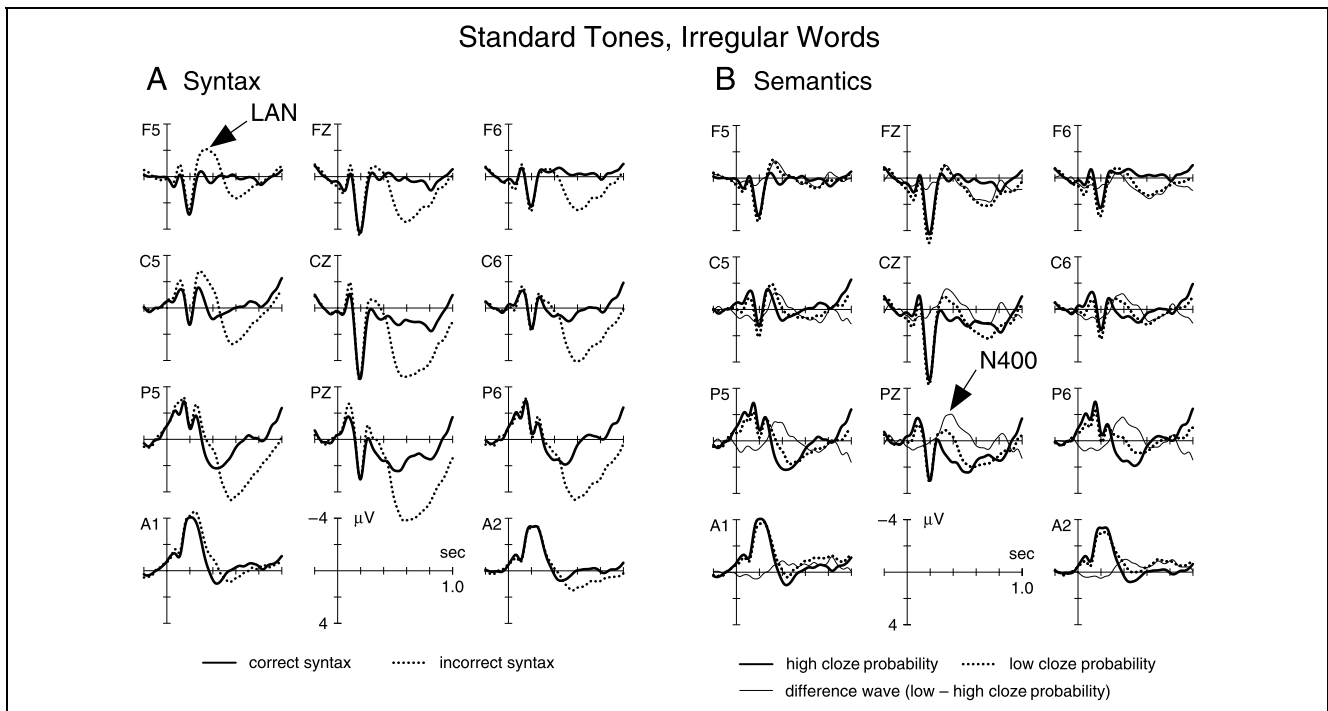


Figure 8. (A) ERPs elicited on standard tones when the last word was syntactically correct (solid line) or syntactically incorrect (dotted line). As in Experiment 1, syntactically incorrect words elicited an LAN (arrow). (B) ERPs elicited by standard tones when the last word had a high (solid line) or low cloze probability (dotted line). Semantically unexpected words elicited an N400 (arrow).

cloze probability [$F(1,21) = 7.21, p < .02$, no two-way interaction, $p > .8$]. The N400 was not lateralized, in contrast to Experiment 1. However, an ANOVA comparing the lateralization of the N400 between both experiments did not indicate a significant difference ($p > .1$).

Syntax × Tones

The former section described the LAN elicited by syntactically incorrect words when sequences ended on standard tones. Figure 9 depicts this LAN effect in the solid difference wave (syntactically correct subtracted from syntactically incorrect words, when tones were standards). The dotted difference wave of Figure 9 shows, again, the effects of processing syntactically incorrect sentences (syntactically correct subtracted from syntactically incorrect words), but now when words are presented on deviant tones. As can be seen in the difference waves, the amplitude of the LAN did not differ when words were presented simultaneously with standard tones compared to when words were presented with deviant tones. An ANOVA for frontal ROIs for the LAN time window (300–450 msec) with factors Syntax and Tone revealed an effect of syntax [$F(1,21) = 5.62, p < .05$], but no interaction between the factors Syntax and Tone [$F(1,21) = 0.04, p > .8$]. It seems that the P600 effect was smaller when elicited on deviant tones than on standard tones, but this difference was statistically not significant ($p > .1$), even when

analyzing only one parietal ROI comprising the electrodes Cz, CP3, CP4, Pz, P3, and P4.

Semantics × Tones

As the LAN, and as in Experiment 1, the N400 elicited by low cloze-probability words virtually was not affected when words were presented together with deviant tones (Figure 10). An ANOVA for parietal ROIs for the N400 time window (300–450 msec) with factors Cloze probability and Tone revealed an effect of cloze probability [$F(1,21) = 4.45, p < .05$], an effect of tone [$F(1,21) = 58.44, p < .0001$], and no interaction between factors Cloze probability and Tone ($p > .1$; the small difference in N400 amplitudes observable in the ERPs of Figure 10 was also not significant when analyzing only one parietal ROI comprising the electrodes Cz, CP3, CP4, Pz, P3, and P4).

Discussion

The data of Experiment 2 show that (language-)syntactic processing does not interact with the processing of physically deviant tones: Compared to the LAN elicited when words were presented simultaneously with standard tones, the amplitude of the LAN was not affected when words were presented simultaneously with deviant tones (although the deviant tones had strong effects on the ERPs). Likewise, and as in Experiment 1, the processing of the semantic aspects of language (as

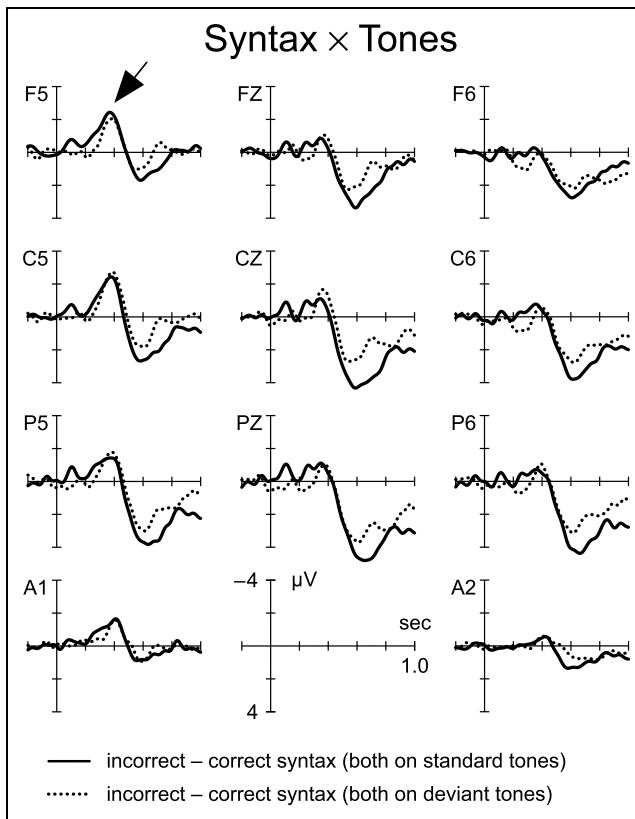


Figure 9. LAN effects (difference waves, syntactically correct subtracted from syntactically incorrect words) for the conditions in which words were presented on standard tones (solid line), and on deviants (dotted line). The amplitude of the LAN did not differ between the two conditions (arrow).

indexed by the N400) was not affected by the processing of the deviant tones.

The MMN was followed by a late negativity (RON), in contrast to Experiment 1, where no late negativity followed the ERAN. It thus appears that the RON is less sensitive to simultaneous processing of tones and linguistic information than the N5. However, it is also possible that the RON simply had larger amplitude values than the N5, and that the RON was, thus, not as strongly compensated by the late positivity (present in a similar time range) than the N5.

GENERAL DISCUSSION

The results of Experiment 1 demonstrate that processing of linguistic syntax (as reflected in the LAN) interacts with the processing of musical syntax (as reflected in the ERAN). Even when the amplitude of the LAN elicited on an irregular chord is compared with the amplitude of the LAN elicited on a regular chord (i.e., even if the LAN could partly overlap with an ERAN, which could lead to an additive effect of LAN and ERAN), the LAN is clearly reduced when elicited during the presentation of an irregular chord. The data of Experiment 2 show that this

interaction is not due to a general effect of deviance-related negativities that precede an LAN: The LAN was not affected when words were presented on physically deviant tones (which elicited an MMN).

These findings are in line with studies suggesting that the neural processes underlying the generation of the ERAN are different from those underlying the generation of the physical MMN (Koelsch, Gunter, Schröger, Tervaniemi, et al., 2001; Maess et al., 2001). Note that previous studies rather suggest a strong overlap of cerebral structures and neural processes involved in the processing of musical syntax with those involved in the processing of linguistic syntax (Patel, 1998, 2003; Tillmann, Janata, et al., 2003; Koelsch, Gunter, von Cramon, et al., 2002; Maess et al., 2001; Patel et al., 1998). With respect to the ERAN, the mentioned study from Maess et al. (2001) suggested that the main neural generators of the ERAN are located in the inferior fronto-lateral cortex (in both hemispheres, with right-hemispheric weighting; see also Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005; Koelsch, Gunter, von Cramon, et al., 2002), areas that are also crucially involved in the processing of linguistic syntax (especially in the left hemisphere; e.g., Friederici, 2002). The finding that language-syntactic processing interacts with music-syntactic processing strongly supports

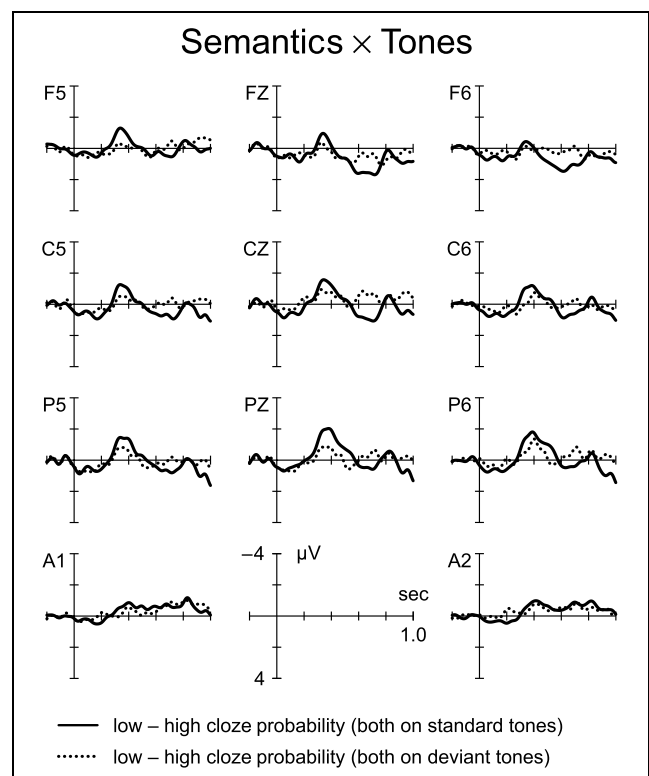


Figure 10. N400 effects (difference waves, high cloze probability subtracted from low cloze probability words) for the conditions in which words were presented on standard tones (solid line) and on deviants (dotted line). The amplitude of the N400 virtually did not differ between the two conditions.

the assumption of such a neural overlap. It is possible that the neural resources for syntactic processing were at least partly consumed by the (quite automatic) processing of the music-syntactic irregularities, resulting in a decrease of resources involved in the generation of the LAN. This finding is surprising, given that the attentional focus of participants was directed on the linguistic information.

With respect to the overlap of neural resources for syntactic processing, the interpretation of the present findings follows the SSIRH (e.g., Patel, 2003), which assumes that the overlap in the syntactic processing of language and music can be conceived of as overlap in the neural areas and operations which provide the resources of syntactic integration. The present results extend this hypothesis in the sense that they indicate that neural resources for syntactic processing are not only shared on the level of syntactic integration (reflected in the P600 from around 600 msec poststimulus on), but already at earlier processing stages (reflected in the present study in the LAN which had an onset at around 250 msec). This earlier stage appears to be important for thematic assignment on the basis of morphosyntactic information during sentence processing (Friederici, 2002). Other even earlier syntactic processing stages comprise initial syntactic structure building: It is assumed that such initial structure building is in the language domain reflected in the ELAN (Friederici, 2002), and it appears that such processes are reflected in the music domain in the ERAN. Future studies could investigate if processing of syntactic information interacts between music and language even at these early stages of syntactic structure building.

Note that on a more abstract level, the processing of both linguistic and musical syntax relies on neural mechanisms that mediate the processing of sequential information, particularly the computation of the relation between a sequential event on the one side, and a context of sequential information that is structured according to complex regularities on the other. These mechanisms appear at least partly to be located in premotor areas (Janata & Grafton, 2003; Huettel, Mack, & McCarthy, 2002; Schubotz & von Cramon, 2001, 2002), comprising the ventro-lateral premotor cortex and BA 44 (in the left hemisphere often referred to as Broca's area). That is, from the view of functional neuroanatomy it is quite plausible that the processing of syntax in music interacts with the processing of syntax in language: The processing of both musical and linguistic syntax requires the activation of neural resources that mediate the processing of complex, regularity-based sequential information.

Interestingly, the scalp topography of the LAN is markedly affected by the type of the accompanying acoustic stimulus: The LAN was stronger lateralized in Experiment 2 (where words were presented on tones) than in Experiment 1. The strong lateralization of the

LAN in Experiment 2 is more characteristic of LAN topographies reported in the literature (e.g., Gunter, Friederici, et al., 2000), whereas the distribution of the LAN in Experiment 1 was much broader. This difference in topographies appears to be related to the more interactive processing of musical and linguistic information in Experiment 1.

Conclusions

The present study investigated neurophysiological correlates of the simultaneous processing of music and language. The ERPs indicate that processing of musical syntax as reflected in the ERAN interacts with processing of linguistic syntax as reflected in the LAN. The processing of physical deviants (indexed by the MMN) did not interact with the processing of linguistic syntax, indicating (a) that ERAN and MMN have different effects on the LAN (underlining the different functional significance of these two deviance-related negativities), and indicating (b) that the interaction between ERAN and LAN is not the result of a general effect of deviance-related negativities on the LAN. Results thus provide direct evidence for shared neural resources engaged for the processing of syntax in language and in music.

The semantic processing of words (indexed by the N400) was not influenced by the processing of the irregular chords. This result was observed under a condition in which participants focused their attention on the sentences, and in which participants made judgments about the semantic and syntactic correctness of words. It is still possible that processing of musical syntax and linguistic semantics interacts under different task conditions, or with different musical stimuli; this issue remains to be specified.

METHODS

Experiment 1

Subjects

Twenty-six right-handed nonmusicians (19–30 years, mean 24.1 years; 15 women) with normal hearing (according to self-report) and normal or corrected-to-normal vision participated in the experiment. Subjects did not have any special musical education (none of them had participated in extracurricular music lessons or performances).

Stimuli

Seventy-eight different chord sequences were used, each chord sequence consisted of five chords (sequences had already been used in some previous studies, e.g., Koelsch, Grossmann, et al., 2003; Koelsch, Schmidt, & Kansok, 2002; Koelsch, Schröger, et al., 2002; Koelsch, Gunter, Schröger, Tervaniemi, et al., 2001; Koelsch,

Gunter, Friederici, et al., 2000). The first chord was the tonic of the following sequence, chords at the second position were: tonic, mediant, submediant, subdominant; at the third position: subdominant, dominant, dominant six–four chord; at the fourth position: dominant seventh chord; at the fifth position: tonic or Neapolitan (sixth) chord. Tonic and Neapolitan chords occurred equiprobably (i.e., 39 sequences ended on a tonic and 39 sequences ended on a Neapolitan chord). Sequences were composed in different voicings (e.g., starting with the root, the third, or the fifth in the soprano voice). Part writing was according to the classical rules of harmony. Stimuli were generated with a piano sound (General MIDI #1) under computerized control via MIDI on a Roland JV-2080 synthesizer (Hamamatsu, Japan). Presentation time of Chords 1–4 was 600 msec, whereas the fifth chord was presented for 1200 msec. There was no silent period between chords or chord sequences; one chord sequence directly followed the other. All final chords had the same loudness and the same decay of loudness, chords were played with approximately 55 dB SPL. Each chord was presented simultaneously with a word (words were presented visually), presentation time was identical for chords and words, and no blank screen was presented between two words.

The sentences were constructed out of 39 sentences that had already been used in the study from Gunter, Friederici, et al. (2000) (the sentence *Er lutscht das Bonbon* was discarded). To all sentences, an adjective was added after the third word (adjectives fitted semantically to the high cloze-probability noun presented at the end of the sentence), so that both sentences and chord sequences had the same number of elements. Each sentence was presented twice during the experiment (once on a tonic and once on a Neapolitan chord). Stimuli were presented in one block of 234 experimental sentences. The order of sentences was pseudorandomized in a way that no correct sentence or its incorrect variation directly followed each other. The ordering of conditions across sentences was balanced. Across the experiment, the stimulation was interrupted for 23 times after a sequence (resulting in 24 sub-blocks), and participants were asked whether the last sentence was correct or incorrect (see also below). After such an inquiry, the tonal key of the following sub-block changed (sequences within one sub-block were in the same key, each of the 12 major keys was used in two sub-blocks, the order of keys was pseudorandomized).

Procedure

Testing was carried out in an acoustically and electrically shielded EEG cabin. Subjects were seated in a comfortable chair facing a computer monitor at a distance of 1.15 m. The task was to ignore the music, to concentrate on the visually presented sentences, and to judge for

each sentence whether the last word was syntactically or semantically correct or incorrect.

Participants were informed that the stimulation would infrequently be interrupted, and that they will be asked then via the monitor whether the sentence preceding the interruption was correct or incorrect (subjects reported their answer by pressing one of two response buttons). Participants were only informed about the different sentence types, not about the Neapolitan chords or their nature. To familiarize participants with the task, two examples of the possible sentence violations were presented, both with sequences ending on a regular (tonic) chord. The experimental session had a duration of approximately 25 minutes.

Recordings and Data Analysis

The EEG was recorded from 60 Ag–AgCl electrodes placed on the head according to the expanded 10–20 system. The reference electrode was placed on the tip of the nose. Sampling rate was 250 Hz (for each channel) and data were filtered with a 70-Hz antialiasing filter during data acquisition. Horizontal and vertical EOGs were recorded bipolarly. Electrode resistance was kept below 5 k Ω .

For elimination of artifacts caused by eye movements, sampling points were rejected off-line whenever the standard deviation within a 200-msec window centered around a sampling point exceeded 35 μ V in the vertical, and 25 μ V in the horizontal EOG. The analogous procedure was carried out for all other electrodes to eliminate artifacts caused by drifts or body movements, with a 500-msec gliding window and a threshold of 25 μ V standard deviation (at any electrode). Averaged waveforms were aligned to a 200-msec prestimulus baseline. For statistical evaluation, ERPs were analyzed by repeated-measures ANOVAs as univariate tests of hypotheses for within-subjects effects (if not separately indicated). Mean ERP values were computed for four ROIs: left anterior (F7, F3, FT7, FC3), right anterior (F4, F8, FC4, FT8), left posterior (P3, P5, CP3, TP7), and right posterior (P4, P6, CP4, TP8). Possible factors that entered the ANOVAs were Cloze probability (high \times low), Syntax (regular \times irregular), Hemisphere (left \times right ROIs), and Chord type (regular [tonic] \times irregular [Neapolitan] chords). Time windows used for statistical analyses were 150–250 msec (ERAN), 300–450 msec (LAN and N400), 350–600 msec (N5), and 450–700 msec (P600/late positive component). After statistical evaluation, grand-averaged ERPs were for presentation purposes filtered with a 10-Hz low-pass filter (41 points, FIR).

Experiment 2

Subjects

Twenty-two right-handed nonmusicians (20–30 years, mean 24.0 years; 10 women) with normal hearing

(according to self-report) and normal or corrected-to-normal vision participated in the experiment. Subjects did not have any special musical education and none of them had participated in extracurricular music lessons or performances. None of the subjects had participated in Experiment 1.

Stimuli

Stimuli were similar to those used in Experiment 1, except that chords were replaced by single tones. Tones at Positions 1 to 4 were standard tones with a frequency of 440 Hz, played with a piano sound (General MIDI #1). Tones at Position 5 were either standard tones ($p = .5$) or physically deviant tones. Four types of physical deviances were used (required due to the equiprobability of standards and deviants at the sequence ending): a frequency deviant (496 Hz), two intensity deviants (one being 60% louder, the other one being 60% softer than the standard loudness), and timbre deviants that had an instrumental timbre different from the standard piano timbre (e.g., marimba, organ, trumpet). As in Experiment 1, stimuli were generated under computerized control via MIDI on the Roland JV-2080 synthesizer.

Procedure, Recordings, and Data Analysis

Procedure, recordings, and data analysis were identical to Experiment 1, except that (a) different time windows were used for the statistical analysis of MMN (90–190 msec), and RON (600–900 msec), and (b) that for statistical comparison of ERPs the factor Chord-type was replaced by the factor Tone (standard \times deviant tones).

Reprint requests should be sent to Stefan Koelsch, Max-Planck-Institute of Cognitive Neuroscience, Stephanstr. 1a, 04103 Leipzig, Germany, or via e-mail: koelsch@cbs.mpg.de.

Notes

1. For details concerning grammatical gender and the German gender system, see Gunter, Friederici, et al. (2000).
2. Gunter, Friederici, et al. (2000) report for their sentences a high cloze probability of 74%, and a low cloze probability of 15%. Syntactically incorrect sentences with low cloze probability used in that study were not used in the present study.
3. With a common average reference, a small N5 was visible in the ERP waveforms (not shown), but again, statistical analysis did not reveal a significant N5 effect.

REFERENCES

Besson, M., Faita, F., Peretz, I., Bonnel, A. M., & Requin, J. (1998). Singing in the brain: Independence of lyrics and tunes. *Psychological Science*, *9*, 494–498.

- Besson, M., & Schön, D. (2001). Comparison between language and music. In R. J. Zatorre & I. Peretz (Eds.), *The biological foundations of music* (pp. 232–258). New York: The New York Academy of Sciences.
- Bonnel, A. M., Faita, F., Peretz, I., & Besson, M. (2001). Divided attention between lyrics and tunes of operatic songs: Evidence for independent processing. *Perception & Psychophysics*, *63*, 1201–1213.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78–84.
- Friederici, A. D. (2004). The neural basis of syntactic processing. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (3rd ed., pp. 289–302). Cambridge: MIT Press.
- Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, *12*, 556–568.
- Gunter, T. C., Stowe, L., & Mulder, G. (1997). When syntax meets semantics. *Psychophysics*, *34*, 660–676.
- Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, *13*, 339–356.
- Huetzel, S. A., Mack, P. B., & McCarthy, G. (2002). Perceiving patterns in random series: Dynamic processing of sequence in prefrontal cortex. *Nature Neuroscience*, *5*, 485–490.
- Janata, P., Birk, J. L., Van Horn, D. J., Leman, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying western music. *Science*, *298*, 2167–2170.
- Janata, P., & Grafton, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*, *6*, 682–687.
- Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, *15*, 207–212.
- Koelsch, S., & Friederici, A. D. (2003). Towards the neural basis of different neurophysiological investigation methods (EEG, MEG, fMRI). *Annals of the New York Academy of Sciences*, *999*, 15–27.
- Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *Neuroimage*, *25*, 1068–1076.
- Koelsch, S., Grossmann, T., Gunter, T. C., Hahne, A., & Friederici, A. D. (2003). Children processing music: Electric brain responses reveal musical competence and gender differences. *Journal of Cognitive Neuroscience*, *15*, 683–693.
- Koelsch, S., Gunter, T. C., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: “Non-musicians” are musical. *Journal of Cognitive Neuroscience*, *12*, 520–541.
- Koelsch, S., Gunter, T. C., Schröger, E., & Friederici, A. D. (2003). Processing tonal modulations: An ERP-study. *Journal of Cognitive Neuroscience*, *15*, 1149–1159.
- Koelsch, S., Gunter, T. C., Schröger, E., Tervaniemi, M., Sammler, D., & Friederici, A. D. (2001). Differentiating ERAN and MMN: An ERP-study. *NeuroReport*, *12*, 1385–1389.
- Koelsch, S., Gunter, T. C., von Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical “language-network” serves the processing of music. *Neuroimage*, *17*, 956–966.
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T. C., & Friederici, A. D. (2004). Music, language, and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*, 302–307.

- Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Influences of musical expertise on the ERAN: An ERP-study. *Psychophysics*, *39*, 657–663.
- Koelsch, S., Schröger, E., & Gunter, T. C. (2002). Music matters: Preattentive musicality of the human brain. *Psychophysics*, *39*, 1–11.
- Kutas, M., & Federmeier, K. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463–470.
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). “Musical syntax” is processed in the area of Broca: An MEG-study. *Nature Neuroscience*, *4*, 540–545.
- Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.
- Parsons, L. (2001). Exploring the functional neuroanatomy of music performance, perception, and comprehension. *Annals of the New York Academy of Sciences*, *930*, 211–231.
- Patel, A. D. (1998). Syntactic processing in language and music: Different cognitive operations, similar neural resources? *Music Perception*, *16*, 27–42.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*, 674–681.
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*, 717–733.
- Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S. V., Lechevalier, B., & Eustache, F. (1997). The structural components of music perception, a functional anatomical study. *Brain*, *120*, 229–243.
- Poulin-Charronnat, B., Bigand, E., Madurell, F., & Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition*, *94*, B67–B78.
- Schröger, E. (1998). Measurement and interpretation of the mismatch negativity (MMN). *Behavior Research Methods, Instruments, & Computers*, *30*, 131–145.
- Schröger, E., Giard, M. H., & Wolff, C. (2000). Auditory distraction: Event-related potential and behavioral indices. *Clinical Neurophysiology*, *111*, 1450–1460.
- Schröger, E., & Wolff, C. (1998). Attentional orienting and reorienting is indicated by human event-related brain potentials. *NeuroReport*, *9*, 3355–3358.
- Schubotz, R. I., & von Cramon, D. Y. (2001). Interval and ordinal properties of sequences are associated with distinct premotor areas. *Cerebral Cortex*, *11*, 210–222.
- Schubotz, R. I., & von Cramon, D. Y. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: An fMRI study. *Neuroimage*, *15*, 787–796.
- Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music*. New York: Oxford University Press.
- Tillmann, B., Bharucha, J., & Bigand, E. (2000). Implicit learning of tonality: A self-organized approach. *Psychological Review*, *107*, 885–913.
- Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research*, *16*, 145–161.
- Zatorre, R., & Peretz, I. (Eds.) (2001). *The biological foundations of music*. New York: NY Academy of Sciences.